|  |  |  |  |
| --- | --- | --- | --- |
| 1 | 00:00:00,678 --> 00:00:01,398 | 大家下午好 |  |
| 2 | 00:00:01,868 --> 00:00:02,756 | 我是王旭 |  |
| 3 | 00:00:03,034 --> 00:00:05,775 | 今天是会向大家汇报 |  |
| 4 | 00:00:06,075 --> 00:00:06,921 | 我们三年以来 |  |
| 5 | 00:00:06,921 --> 00:00:08,246 | 取得的一个成果 |  |
| 6 | 00:00:08,587 --> 00:00:10,790 | 就是基于RISC-V指令集的 |  |
| 7 | 00:00:10,790 --> 00:00:12,125 | Egret的系列处理器 |  |
| 8 | 00:00:13,778 --> 00:00:15,121 | 今天的报告会以 |  |
| 9 | 00:00:15,121 --> 00:00:16,496 | 以下四个方面展开 |  |
| 10 | 00:00:16,996 --> 00:00:19,799 | 会详细的介绍Egret系列处理器的 |  |
| 11 | 00:00:20,260 --> 00:00:21,780 | 特性和应用场景 |  |
| 12 | 00:00:23,593 --> 00:00:24,593 | 首先介绍一下 |  |
| 13 | 00:00:24,593 --> 00:00:25,787 | 我们的基本情况 |  |
| 14 | 00:00:27,618 --> 00:00:29,246 | 这一个项目是由 |  |
| 15 | 00:00:29,246 --> 00:00:30,760 | 厦门半导体投资集团 |  |
| 16 | 00:00:30,780 --> 00:00:32,565 | 和清华大学集成电路学院 |  |
| 17 | 00:00:32,565 --> 00:00:33,706 | 联合进行开发的 |  |
| 18 | 00:00:34,100 --> 00:00:36,700 | 用三年的时间我们积累了两款 |  |
| 19 | 00:00:37,105 --> 00:00:39,365 | 单核的CPU和一款多核的CPU |  |
| 20 | 00:00:39,405 --> 00:00:41,703 | 以及一款神经网络加速器 |  |
| 21 | 00:00:42,584 --> 00:00:44,365 | 那么CPU的性能 |  |
| 22 | 00:00:44,365 --> 00:00:47,181 | 是能够和ARM Cortex-A7进行对标的 |  |
| 23 | 00:00:47,831 --> 00:00:50,870 | 这一款神经网络加速器 |  |
| 24 | 00:00:50,871 --> 00:00:53,153 | 它的性能能够达到2TOPS |  |
| 25 | 00:00:53,550 --> 00:00:54,940 | 这些成果的取得 |  |
| 26 | 00:00:54,940 --> 00:00:56,650 | 离不开团队成员 |  |
| 27 | 00:00:56,930 --> 00:00:59,130 | 多年以来在处理器设计和验证 |  |
| 28 | 00:00:59,505 --> 00:01:00,631 | 相关积累的经验 |  |
| 29 | 00:01:02,025 --> 00:01:03,409 | 我们这个团队 |  |
| 30 | 00:01:03,409 --> 00:01:05,909 | 是由清华大学的何虎老师带队 |  |
| 31 | 00:01:06,290 --> 00:01:08,484 | 核心的研发人员 |  |
| 32 | 00:01:08,750 --> 00:01:11,143 | 均是毕业于清华大学集成电路学院 |  |
| 33 | 00:01:11,603 --> 00:01:14,428 | 我们这个团队的人员虽然不多 |  |
| 34 | 00:01:14,655 --> 00:01:16,178 | 但是很完整 |  |
| 35 | 00:01:16,531 --> 00:01:18,281 | 覆盖了CPU的设计 |  |
| 36 | 00:01:18,615 --> 00:01:19,615 | CPU的验证 |  |
| 37 | 00:01:19,840 --> 00:01:21,190 | NPU的开发 |  |
| 38 | 00:01:21,190 --> 00:01:22,856 | 以及软件工具链的支持 |  |
| 39 | 00:01:24,106 --> 00:01:26,790 | 做CPU是需要一个长时间 |  |
| 40 | 00:01:27,081 --> 00:01:29,153 | 积累迭代的过程 |  |
| 41 | 00:01:29,153 --> 00:01:30,580 | 才能做好的一件事情 |  |
| 42 | 00:01:30,871 --> 00:01:33,437 | 我们前期的积累的一个 |  |
| 43 | 00:01:33,437 --> 00:01:34,543 | 典型的成果 |  |
| 44 | 00:01:34,543 --> 00:01:36,943 | 是一个基于自研指令集的 |  |
| 45 | 00:01:37,328 --> 00:01:38,162 | 超长指令字的 |  |
| 46 | 00:01:38,162 --> 00:01:39,803 | 一个密码处理器IP |  |
| 47 | 00:01:40,193 --> 00:01:42,596 | 这款IP是一个 |  |
| 48 | 00:01:43,056 --> 00:01:45,187 | 异构多核的DSP |  |
| 49 | 00:01:45,365 --> 00:01:47,600 | 在几年前已经达到了 |  |
| 50 | 00:01:47,600 --> 00:01:49,568 | 大规模量产这样一个程度 |  |
| 51 | 00:01:50,343 --> 00:01:51,400 | 并且荣获了 |  |
| 52 | 00:01:51,660 --> 00:01:54,800 | 北京市2018年度的科学技术一等奖 |  |
| 53 | 00:01:54,840 --> 00:01:56,171 | 也是一个很高的殊荣 |  |
| 54 | 00:01:58,193 --> 00:02:01,556 | 我们流片的另一个芯片是 |  |
| 55 | 00:02:01,903 --> 00:02:04,490 | 基于ARM指令集的一个通用DSP |  |
| 56 | 00:02:04,818 --> 00:02:06,731 | 它采用的是架构的创新 |  |
| 57 | 00:02:06,731 --> 00:02:08,856 | 兼顾了很好的通用性 |  |
| 58 | 00:02:08,856 --> 00:02:09,975 | 和数据处理能力 |  |
| 59 | 00:02:11,043 --> 00:02:13,921 | 这款芯片也是成功的进行了流片的验证 |  |
| 60 | 00:02:15,459 --> 00:02:17,418 | 在前期这些CPU |  |
| 61 | 00:02:17,418 --> 00:02:18,965 | 相关经验的积累下 |  |
| 62 | 00:02:19,340 --> 00:02:21,960 | 我们在2018年开始 |  |
| 63 | 00:02:21,960 --> 00:02:23,781 | 用了三年的时间致力于 |  |
| 64 | 00:02:23,831 --> 00:02:26,031 | 基于RISC-V指令集的开发 |  |
| 65 | 00:02:27,706 --> 00:02:31,721 | 我们积累了三个系列CPU的IP |  |
| 66 | 00:02:32,021 --> 00:02:35,271 | 分别是B系列 V系列 D系列 |  |
| 67 | 00:02:35,471 --> 00:02:36,928 | 还有一款神经网络加速器 |  |
| 68 | 00:02:37,725 --> 00:02:40,661 | 那么从各个系列来看 |  |
| 69 | 00:02:40,661 --> 00:02:42,262 | 它是一个迭代的过程 |  |
| 70 | 00:02:42,987 --> 00:02:46,715 | 前面两款单核的已经进行了流片 |  |
| 71 | 00:02:46,965 --> 00:02:49,300 | 采用的工艺是SMIC 40nm |  |
| 72 | 00:02:50,309 --> 00:02:51,731 | 这款多核的 |  |
| 73 | 00:02:52,060 --> 00:02:54,496 | 已经是按照今年七月份在UMC |  |
| 74 | 00:02:54,834 --> 00:02:56,520 | 采用28nm工艺流片 |  |
| 75 | 00:02:56,520 --> 00:02:59,390 | 做好了设计验证以及后端 |  |
| 76 | 00:02:59,781 --> 00:03:01,709 | 但是很遗憾 |  |
| 77 | 00:03:02,034 --> 00:03:03,484 | UMC它取消了 |  |
| 78 | 00:03:03,484 --> 00:03:05,409 | 7月份这次MPW的班车 |  |
| 79 | 00:03:05,500 --> 00:03:07,327 | 所以我们只能等到12月份 |  |
| 80 | 00:03:07,637 --> 00:03:10,281 | 再进行下一个时间节点的流片 |  |
| 81 | 00:03:13,103 --> 00:03:14,860 | 这个B系列处理器 |  |
| 82 | 00:03:14,860 --> 00:03:17,086 | 是我们从无到有的 |  |
| 83 | 00:03:17,086 --> 00:03:18,400 | 去设计了这样一款 |  |
| 84 | 00:03:18,803 --> 00:03:20,420 | 自研架构 |  |
| 85 | 00:03:20,421 --> 00:03:21,328 | 顺序双发射 |  |
| 86 | 00:03:21,328 --> 00:03:23,515 | 9级流水线这样一款CPU |  |
| 87 | 00:03:24,246 --> 00:03:27,575 | 它能够支持的是LSU的乱序 |  |
| 88 | 00:03:27,618 --> 00:03:30,037 | 32kb的I-Cache和D-Cache |  |
| 89 | 00:03:30,355 --> 00:03:31,715 | 它最大的特点是能够 |  |
| 90 | 00:03:31,715 --> 00:03:33,631 | pin-to-pin的兼容ARM Cortex-A7 |  |
| 91 | 00:03:34,910 --> 00:03:36,565 | 这款芯片是在 |  |
| 92 | 00:03:37,050 --> 00:03:39,259 | 2019年的12月份进行的流片 |  |
| 93 | 00:03:39,493 --> 00:03:41,240 | 采用的是SMIC 40nm工艺 |  |
| 94 | 00:03:41,681 --> 00:03:43,210 | 经过样片的实测 |  |
| 95 | 00:03:43,211 --> 00:03:44,803 | 面积是1.7平方毫米 |  |
| 96 | 00:03:45,090 --> 00:03:47,065 | 频率能够达到600MHz |  |
| 97 | 00:03:49,106 --> 00:03:50,380 | 在B系列的基础上 |  |
| 98 | 00:03:50,381 --> 00:03:52,400 | 我们又完善了相应的功能 |  |
| 99 | 00:03:52,660 --> 00:03:55,020 | 然后开发了V系列的处理器 |  |
| 100 | 00:03:55,446 --> 00:03:57,696 | 同时向清华大学计算机系 |  |
| 101 | 00:03:57,696 --> 00:03:59,468 | 授权了这一款CPU的IP |  |
| 102 | 00:04:00,050 --> 00:04:02,006 | 将他们自研的向量处理器 |  |
| 103 | 00:04:02,006 --> 00:04:05,053 | 作为协处理器共同的进行开发 |  |
| 104 | 00:04:06,321 --> 00:04:08,628 | 这款处理器是在去年的 |  |
| 105 | 00:04:08,628 --> 00:04:09,993 | 12月份进行的流片 |  |
| 106 | 00:04:09,993 --> 00:04:12,290 | 同样也是采用SMIC 40nm的工艺 |  |
| 107 | 00:04:12,971 --> 00:04:15,271 | 由于现在产能非常紧 |  |
| 108 | 00:04:15,562 --> 00:04:16,556 | 周期变长 |  |
| 109 | 00:04:16,775 --> 00:04:18,475 | 目前这款芯片是正在 |  |
| 110 | 00:04:18,475 --> 00:04:19,225 | 进行封装 |  |
| 111 | 00:04:19,381 --> 00:04:20,995 | 很快会拿到样片 |  |
| 112 | 00:04:21,309 --> 00:04:23,393 | 我们进行更详细的测试 |  |
| 113 | 00:04:26,475 --> 00:04:28,053 | 这款多核的处理器 |  |
| 114 | 00:04:28,090 --> 00:04:30,631 | 是增加了双精度浮点的支持 |  |
| 115 | 00:04:31,080 --> 00:04:33,775 | 并且集成了神经网络加速器 |  |
| 116 | 00:04:34,378 --> 00:04:37,490 | 我们是按照四核进行的开发和验证的 |  |
| 117 | 00:04:37,750 --> 00:04:39,178 | 只不过由于面积的限制 |  |
| 118 | 00:04:39,178 --> 00:04:41,775 | 我们最后是按照双核去流的片 |  |
| 119 | 00:04:41,931 --> 00:04:44,978 | 具备的是512 kb的L2-Cache |  |
| 120 | 00:04:45,834 --> 00:04:48,862 | 它和神经网络加速器是一个 |  |
| 121 | 00:04:49,190 --> 00:04:52,231 | 合作配合去集成在SoC当中 |  |
| 122 | 00:04:54,953 --> 00:04:56,896 | 双精度浮点和单精度浮点 |  |
| 123 | 00:04:56,896 --> 00:04:58,110 | 我们是都支持的 |  |
| 124 | 00:04:58,409 --> 00:05:00,484 | 这是一个我们自研的 |  |
| 125 | 00:05:00,730 --> 00:05:03,770 | FPU的其中一个运算模块 |  |
| 126 | 00:05:03,771 --> 00:05:06,012 | 它支持的是RISC-V定义的 |  |
| 127 | 00:05:06,310 --> 00:05:07,534 | D扩展和F扩展 |  |
| 128 | 00:05:07,878 --> 00:05:10,168 | 同时也支持SIMD的运算 |  |
| 129 | 00:05:12,230 --> 00:05:13,700 | 这一款神经网络加速器 |  |
| 130 | 00:05:13,700 --> 00:05:15,375 | 是一个4核的NPU |  |
| 131 | 00:05:15,500 --> 00:05:18,262 | 它能够支持的是VGG16 |  |
| 132 | 00:05:18,478 --> 00:05:20,356 | ResNet50等典型的网络 |  |
| 133 | 00:05:20,893 --> 00:05:23,540 | 它是以软硬件协作的方式 |  |
| 134 | 00:05:23,540 --> 00:05:25,706 | 来执行CNN的算法 |  |
| 135 | 00:05:26,175 --> 00:05:28,555 | 工作频率在1GHz的时候 |  |
| 136 | 00:05:28,556 --> 00:05:30,435 | 峰值性能能够达到2TOPS |  |
| 137 | 00:05:32,653 --> 00:05:34,403 | 以上就是三款CPU |  |
| 138 | 00:05:34,403 --> 00:05:36,487 | 和一款神经网络加速器的介绍 |  |
| 139 | 00:05:36,753 --> 00:05:38,300 | 下面来介绍一下 |  |
| 140 | 00:05:38,300 --> 00:05:40,337 | Egret的系列处理器的特性 |  |
| 141 | 00:05:41,875 --> 00:05:43,953 | Egret系列处理器 |  |
| 142 | 00:05:43,953 --> 00:05:46,659 | 它能够对标的是ARM Cortex-A7 |  |
| 143 | 00:05:46,889 --> 00:05:49,118 | 能够对它进行原位的替换 |  |
| 144 | 00:05:49,753 --> 00:05:51,434 | 并且迭代了三个版本 |  |
| 145 | 00:05:51,856 --> 00:05:54,230 | 复杂度是逐渐提高的过程 |  |
| 146 | 00:05:55,034 --> 00:05:56,765 | 同时我们还支持 |  |
| 147 | 00:05:57,231 --> 00:05:59,853 | 向量加速器和神经网络加速器 |  |
| 148 | 00:06:00,387 --> 00:06:01,865 | 流片的工艺也是从 |  |
| 149 | 00:06:01,865 --> 00:06:04,353 | 40nm到28nm逐步提升的 |  |
| 150 | 00:06:06,384 --> 00:06:08,353 | UVM验证平台其实对于 |  |
| 151 | 00:06:08,353 --> 00:06:09,580 | CPU的开发来讲 |  |
| 152 | 00:06:09,580 --> 00:06:10,640 | 是至关重要的 |  |
| 153 | 00:06:10,781 --> 00:06:13,362 | 我们团队也是自己搭建了完备的 |  |
| 154 | 00:06:13,620 --> 00:06:15,181 | 单核和多核的验证系统 |  |
| 155 | 00:06:15,640 --> 00:06:18,240 | 能够支持指令精度的动态实时对比 |  |
| 156 | 00:06:18,640 --> 00:06:19,915 | 并且具备成熟的 |  |
| 157 | 00:06:19,915 --> 00:06:22,225 | 多核一致性实时动态验证机制 |  |
| 158 | 00:06:24,750 --> 00:06:26,825 | Egret系列处理器的开发过程中 |  |
| 159 | 00:06:26,825 --> 00:06:28,330 | 涵盖了CPU的设计 |  |
| 160 | 00:06:28,331 --> 00:06:29,115 | UVM验证 |  |
| 161 | 00:06:29,500 --> 00:06:30,481 | SoC集成 |  |
| 162 | 00:06:30,481 --> 00:06:32,937 | 和FPGA原型验证这四部分 |  |
| 163 | 00:06:33,403 --> 00:06:35,209 | 从各个维度上保证了 |  |
| 164 | 00:06:35,209 --> 00:06:36,196 | 功能的可靠性 |  |
| 165 | 00:06:38,680 --> 00:06:40,284 | Egret系列处理器的应用场景 |  |
| 166 | 00:06:40,284 --> 00:06:41,380 | 也是比较广泛的 |  |
| 167 | 00:06:42,690 --> 00:06:44,643 | 它可以应用到汽车 电子 |  |
| 168 | 00:06:44,643 --> 00:06:45,531 | 工业控制 |  |
| 169 | 00:06:45,690 --> 00:06:47,412 | 人工智能 物联网 |  |
| 170 | 00:06:47,412 --> 00:06:48,710 | 安全等多个领域 |  |
| 171 | 00:06:49,231 --> 00:06:52,953 | 我们目前主要攻关的一个方向是 |  |
| 172 | 00:06:52,953 --> 00:06:55,353 | 车用MCU也就是在传动系统 |  |
| 173 | 00:06:55,620 --> 00:06:57,520 | 在汽车的发动机控制 |  |
| 174 | 00:06:57,520 --> 00:07:00,084 | 底盘控制相关的领域做一个应用 |  |
| 175 | 00:07:00,460 --> 00:07:03,106 | 结合CPU和我们自研的NPU |  |
| 176 | 00:07:03,340 --> 00:07:04,800 | 也可以在图像处理 |  |
| 177 | 00:07:04,801 --> 00:07:07,231 | 目标检测领域有所作为 |  |
| 178 | 00:07:07,660 --> 00:07:09,280 | 在其他的应用领域 |  |
| 179 | 00:07:09,281 --> 00:07:11,290 | 我们也可以针对特定的 |  |
| 180 | 00:07:11,290 --> 00:07:12,940 | 应用场景的需求 |  |
| 181 | 00:07:12,941 --> 00:07:14,040 | 进行定制化的开发 |  |
| 182 | 00:07:15,975 --> 00:07:17,210 | 我们的优势在于 |  |
| 183 | 00:07:17,211 --> 00:07:19,284 | 目前我们的团队的配合 |  |
| 184 | 00:07:19,631 --> 00:07:21,015 | 是非常默契的 |  |
| 185 | 00:07:21,593 --> 00:07:24,070 | 涵盖了CPU和NPU的设计 验证 |  |
| 186 | 00:07:24,353 --> 00:07:26,093 | FPGA验证等全流程的环节 |  |
| 187 | 00:07:26,534 --> 00:07:28,900 | RTL全部是我们自行编写的 |  |
| 188 | 00:07:29,765 --> 00:07:31,015 | 三年时间积累了 |  |
| 189 | 00:07:31,015 --> 00:07:32,334 | 三款CPU和一款 |  |
| 190 | 00:07:32,731 --> 00:07:33,690 | 神经网络加速器 |  |
| 191 | 00:07:34,025 --> 00:07:35,650 | 是由单核向多核递进 |  |
| 192 | 00:07:35,965 --> 00:07:37,965 | 并且不断扩展功能 |  |
| 193 | 00:07:38,759 --> 00:07:40,325 | 支持AI加速器的扩展 |  |
| 194 | 00:07:40,934 --> 00:07:42,153 | 同时也采用了 |  |
| 195 | 00:07:42,153 --> 00:07:43,840 | 高性能处理器的设计技术 |  |
| 196 | 00:07:44,178 --> 00:07:45,159 | 支持单精度 |  |
| 197 | 00:07:45,159 --> 00:07:46,128 | 双精度浮点 |  |
| 198 | 00:07:46,415 --> 00:07:47,841 | LSU的乱序执行 |  |
| 199 | 00:07:48,062 --> 00:07:49,606 | 可以支持8核的互联 |  |
| 200 | 00:07:49,606 --> 00:07:50,618 | 和TAGE分支预测 |  |
| 201 | 00:07:51,221 --> 00:07:52,671 | 完备的验证环境 |  |
| 202 | 00:07:53,140 --> 00:07:55,212 | 给我们的设计也提供了强烈的支撑 |  |
| 203 | 00:07:55,600 --> 00:07:57,160 | 单核和多核的验证平台 |  |
| 204 | 00:07:57,160 --> 00:07:59,315 | 能够实现指令精度的实时对比 |  |
| 205 | 00:07:59,787 --> 00:08:01,250 | 也支持多核一致性的 |  |
| 206 | 00:08:01,250 --> 00:08:02,293 | 动态实时验证 |  |
| 207 | 00:08:03,703 --> 00:08:05,462 | 目前我们三年的 |  |
| 208 | 00:08:05,462 --> 00:08:06,990 | 孵化期是即将结束了 |  |
| 209 | 00:08:07,250 --> 00:08:09,193 | 也正在寻求新的合作伙伴 |  |
| 210 | 00:08:09,671 --> 00:08:12,110 | 如果有任何关于RISC-V的问题 |  |
| 211 | 00:08:12,168 --> 00:08:13,184 | 可以扫码 |  |
| 212 | 00:08:13,184 --> 00:08:14,906 | 然后咱们线下进行交流 |  |
| 213 | 00:08:15,612 --> 00:08:17,195 | 我的报告就到这里 |  |
| 214 | 00:08:17,195 --> 00:08:17,650 | 谢谢大家 |  |